1. **Executive Summary**

   Statistical procedures were used to obtain forecasts of United States milk production for each month of 1976, based on monthly data for the years 1962 to 1975. The forecasts show that it is likely that there will be a small increase of production. This increase is forecast to be in the range of 3% to 4%. The forecasts do contain a certain amount of uncertainly, and it is possible that there could be no increase in production, or that the increase could be as large as 8%.

   Detailed Desciption

   Figure 1 below shows a plot of the milk production data.

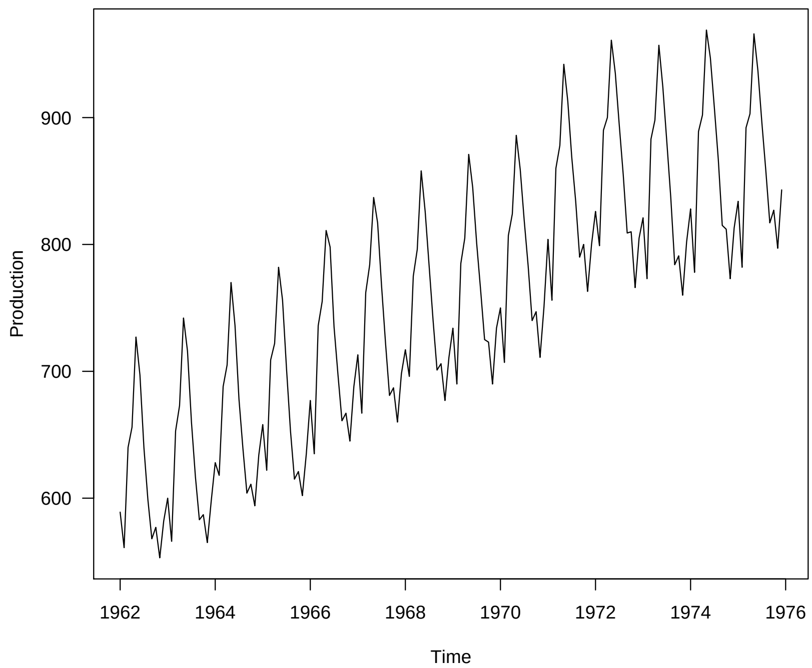

Figure 1: United States milk production.

The plot shows a strong seasonal effect and an upward trend (which levels off towards the end of the series). There is no evidence that the variability of the series is increasing with its mean level, so no variance stabilizing transformation is required.

The first stage of the analysis is to find the amount of differencing requited to make the series stationary. The most visible feature of the series is the seasonal pattern, so the appropriate first step is to apply seasonal differencing. The result is shown in figure 2. The plot shows long-term variation which could represent nonstationarity. This can be checked by looking for slow decay of the autocorrelation function of the differenced series. Figure 3 shows that the ACF does decay slowly and that additional differencing is required. The ACF does not show a seasonal pattern so it is appropriate to use simple differencing.

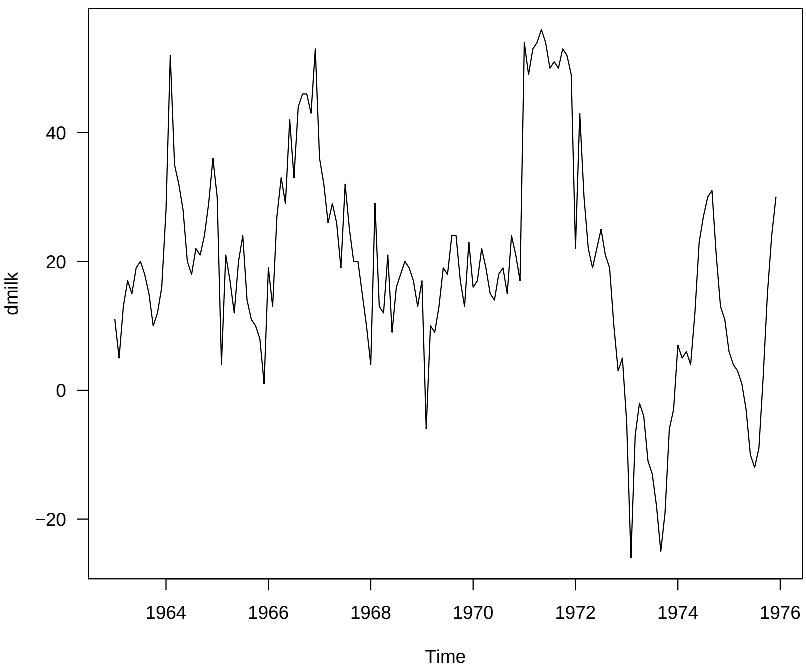

Figure 2: The seasonally differenced milk production series.

The result of applying the two kinds of differencing is shown in figure 4. The series now looks stationary. (This could be confirmed by looking at the ACF shown in figure 5).

An initial model for forecasting can now be determined by examining the ACF and PACF for the twice-differenced series. Plots of the ACF and PACF are shown in figures 5 and 6. The ACF plot shows sharp cutoff; after lag 1 for a nonseasonal component and after lag 12 for the the nonseasonal part. (The large value at lag 13 is a natural consequence of using a product seasonal model). The PACF shows exponential decay at multiples of the seasonal period and what could be seen as cosinusoidal behaviour at non seasonal lags. This suggests using a product moving average model. The appropriate model would seen to

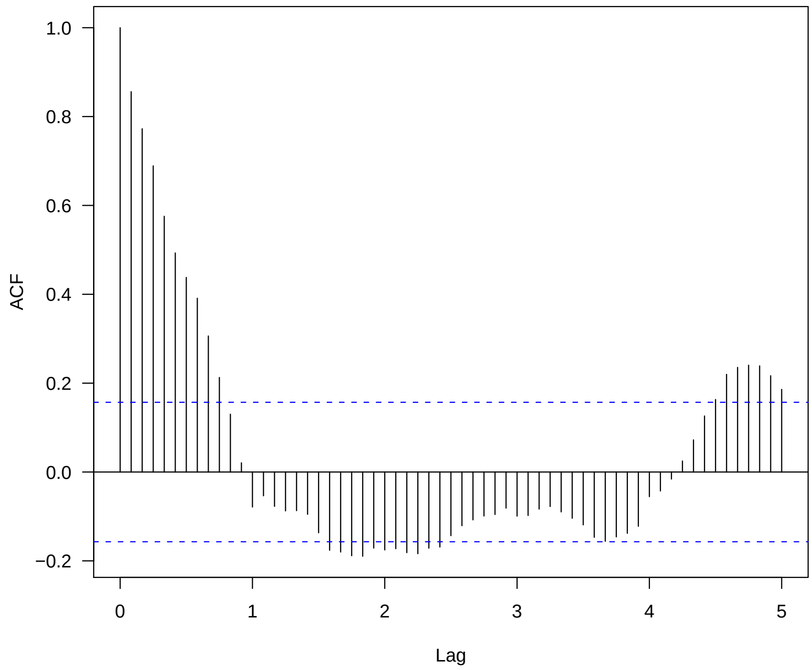

Figure 3: The ACF of the seasonaly differenced milk production series.

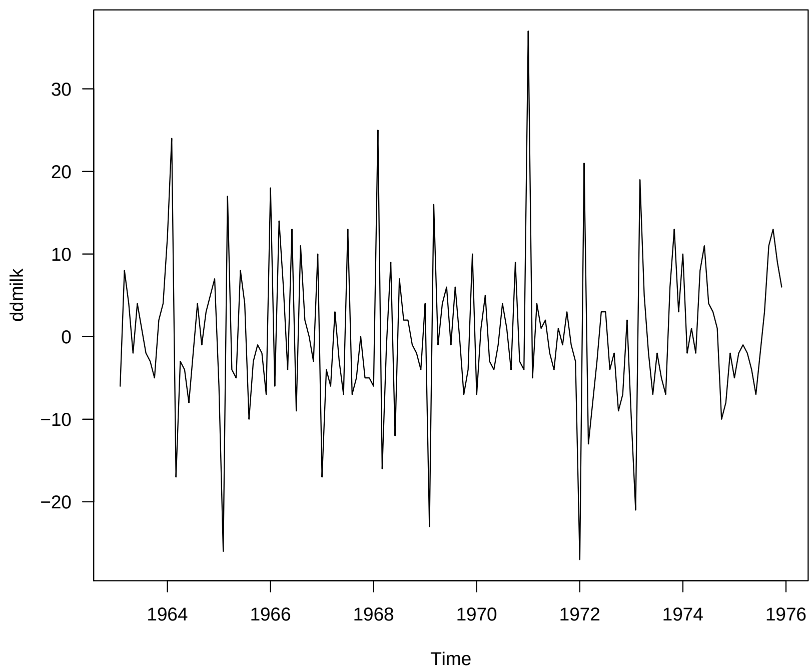

Figure 4: The twice-differenced milk production series.

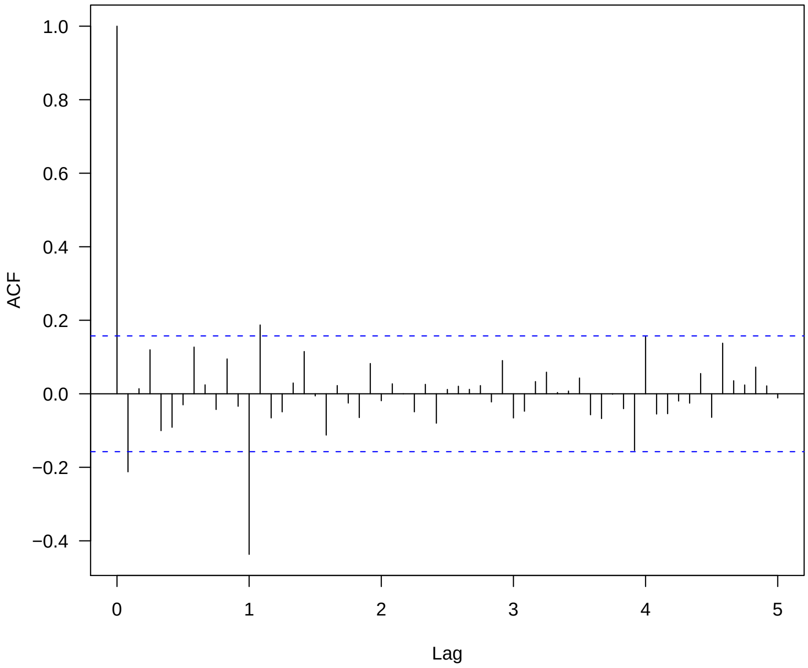

Figure 5: The ACF of the twice-differenced milk production series.

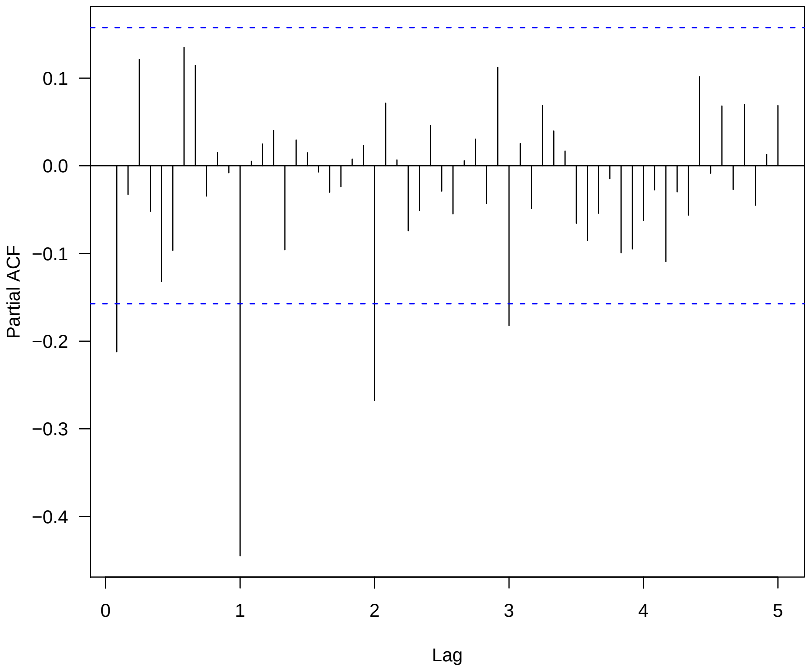

Figure 6: The PACF of the twice-differenced milk production series.

be ARIMA$(0,1,1) \times (0,1,1)_{12}$.

$$\nabla\nabla_{12} Y_t = (1 + \theta_1 L)(1 + \Theta_1 L^{12})\varepsilon_t$$

Fitting this model produces the following results.

	θ_1	Θ_{12}
Coefficient	-0.2204	-0.6214
Standard Error	0.0748	0.0627

Both coefficients are significant, and when additional MA coefficents (simple or seasonal) are added to the model they are not significant. This model is a suitable candidate for making forecasts.

Before making forecasts we will carry out a check of the model residuals. Diagnostic plots are shown in figure 7. They show no problems with the residuals and that we can be confident that the forecasts and their standard errors are reasonable.

Forecasts for United States milk production for the year 1976 are shown in table 1 below and plotted in figure 8. The forecasts show a predicted increase of roughly 3% to 4% in milk production. Because the lower confidence limit for the predictions is close to the values for 1975, we can be fairly sure that there will be small increase in production for 1976.

Table 1: United States Milk Production Forecasts For 1976.

Month	Forecast	Std. Error
Jan	865.0	7.3
Feb	817.7	9.2
Mar	924.4	10.8
Apr	937.5	12.2
May	1000.6	13.4
Jun	973.2	14.6
Jul	931.9	15.6
Aug	892.3	16.6
Sep	846.4	17.6
Oct	851.5	18.5
Nov	817.5	19.3
Dec	859.8	20.1

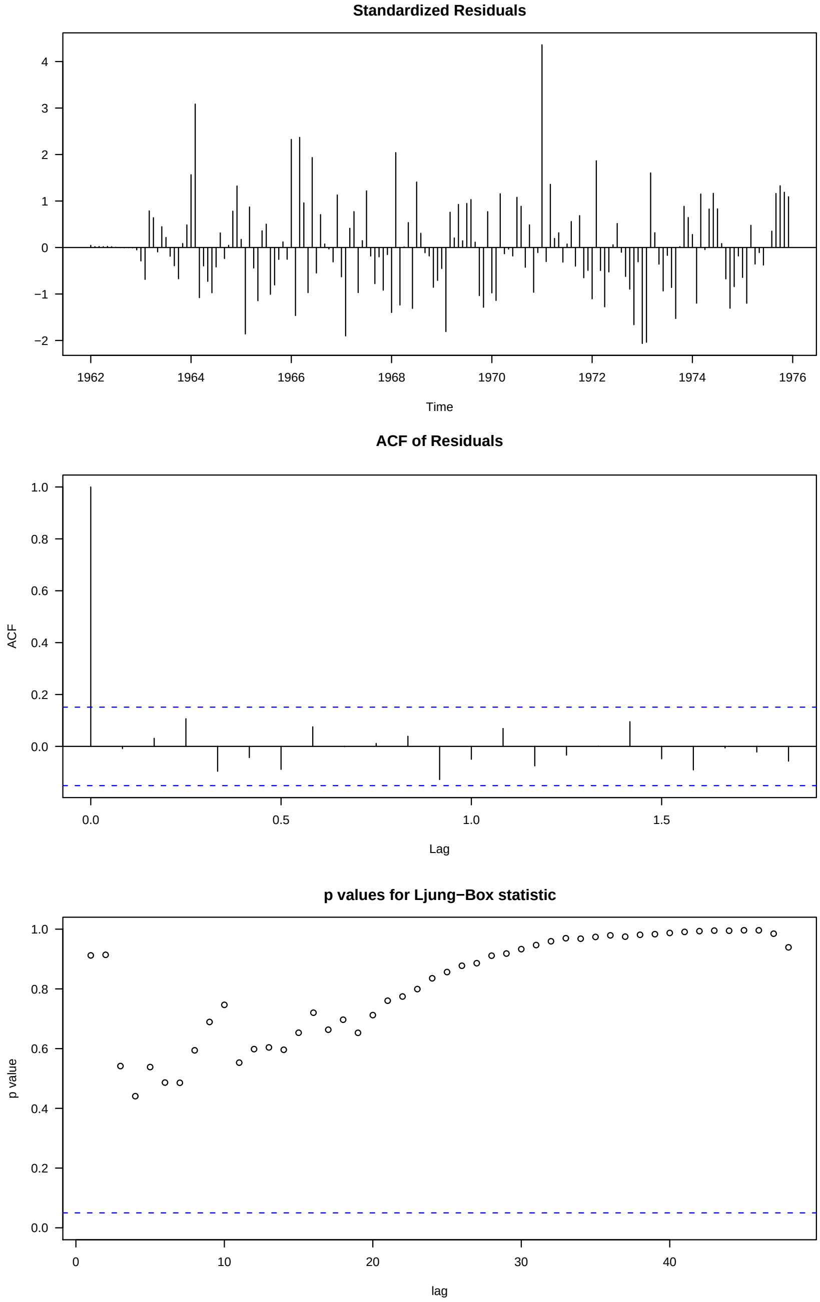

Figure 7: Residual diagnostics for the milk production series.

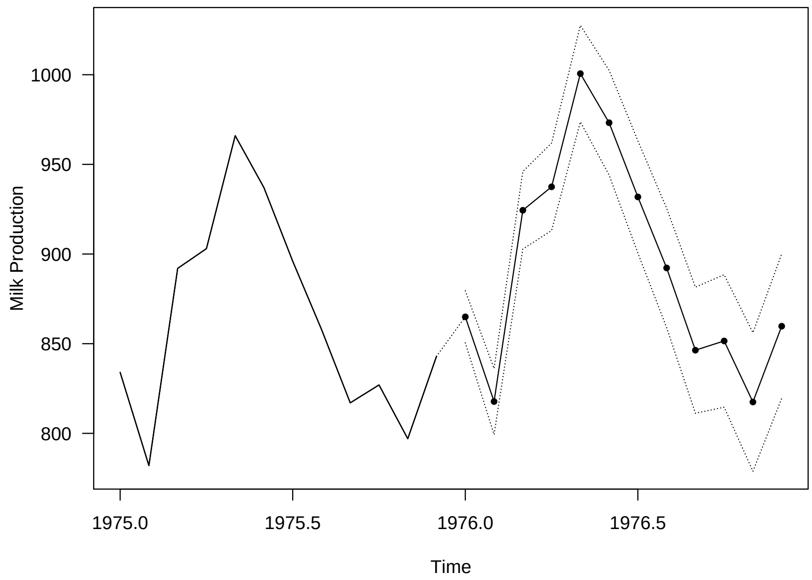

Figure 8: Forecasts for United States milk production.

2. **Executive Summary**

Statistical techniques were used to obtain forecasts for the monthly number of New Zealand road deaths. The study uncovered a basic seasonal pattern which was used to produce forecasts, but the randomness present in the record makes these forecasts unreliable.

Detailed Desciption

Figure 9 below shows a plot of the road death series.

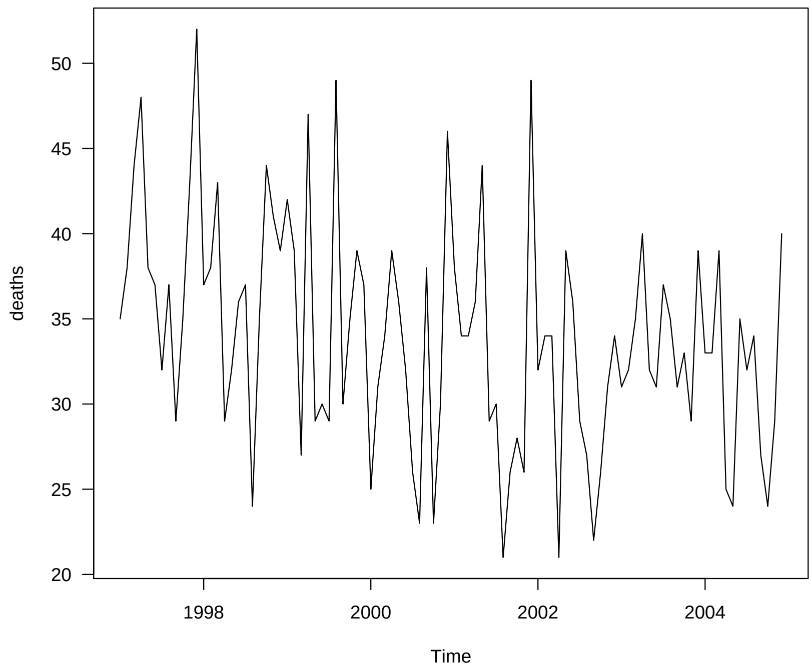

Figure 9: Monthly New Zealand road deaths 1997 to 2004.

The series appears to be stationary so we can examine the ACF and PACF to try to determine an appropriate model. The graphs of the ACF and PACF are shown in figures 10 and 11.

The graphs give very little indication of what an appropriate model might be. Because of this it was decided to start with an ARIMA$(1,0,1)\times(1,0,1)_{12}$ model and to progressively remove terms from the model if they we not significant. The fitted coefficients to the model were.

	ϕ_1	θ_1	Φ_1	Θ_1
Estimate	0.9745	-0.9163	0.9443	-0.8581
Std. Error	0.0347	0.0569	0.1683	0.2600

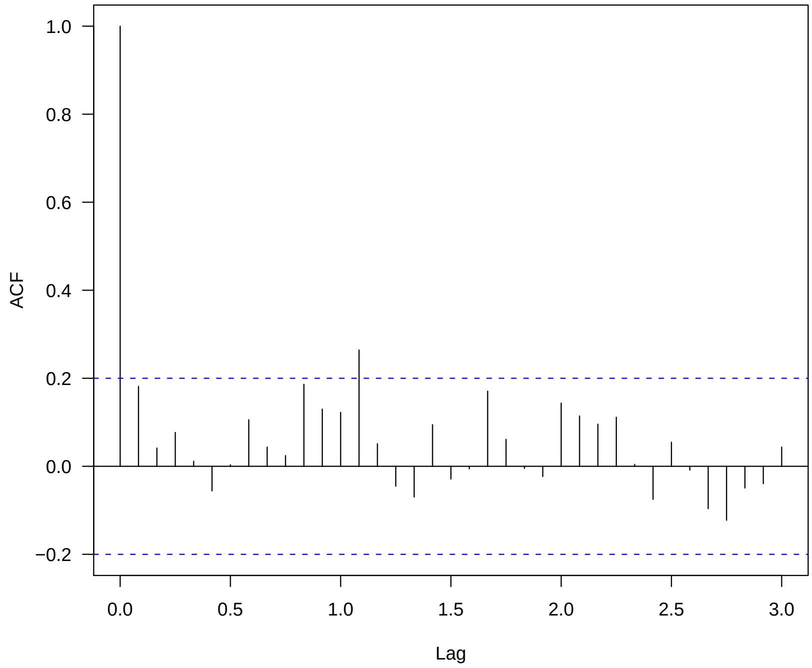

Figure 10: The ACF of the road deaths series.

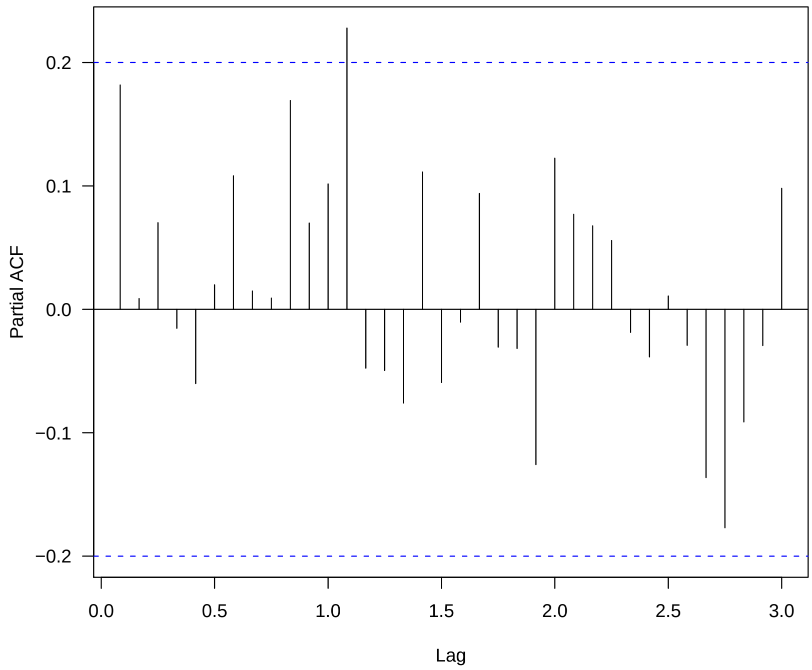

Figure 11: The PACF of the road deaths series.

This is a surprise because *all* of the coefficients are significant. Looking more closely reveals that the simple AR coefficent is close to 1, suggesting that the series is close to nonstationary. This in turn suggest that it might be better to replace the model above by ARIMA$(0,1,1) \times (1,0,1)_{12}$. Fitting this model produces the coefficients

	θ_1	Φ_1	Θ_1
Estimate	-0.9384	0.9980	-0.9722
Std. Error	0.0314	0.0178	0.1243

(There are problems fitting this model. The number of iterations allowed in the optimisation must be raised to get to convergence.

Again, this model seems to indicate that the model is close to nonstationarity because Φ_1 is nearly 1. This suggests that the model ARIMA$(0,1,1)\times(0,1,1)_{12}$ might be appropriate. Fitting this model produces the coefficients

	θ_1	Θ_1
Estimate	-0.9649	-1.000
Std. Error	0.0634	0.221

(This model also has a very low AIC value, which is interesting, but not sufficient reason in itself to choose this model.)

Looking back at the ACF and PACF for the twice differenced data makes it clear that this is a very good model to use. This is confirmed by an inspection of the residual plots.

The forecasts produced from this model are given in table 2 and plotted in figure 14. The width of the confidence intervals around the forecasts means that these forecasts cannot be treated as reliable.

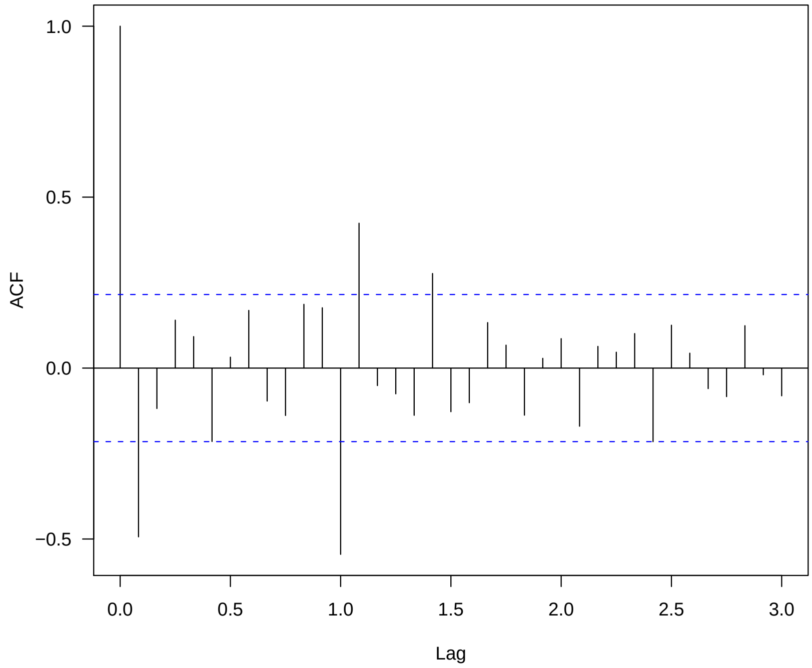

Figure 12: The ACF of the twice-differenced road deaths series.

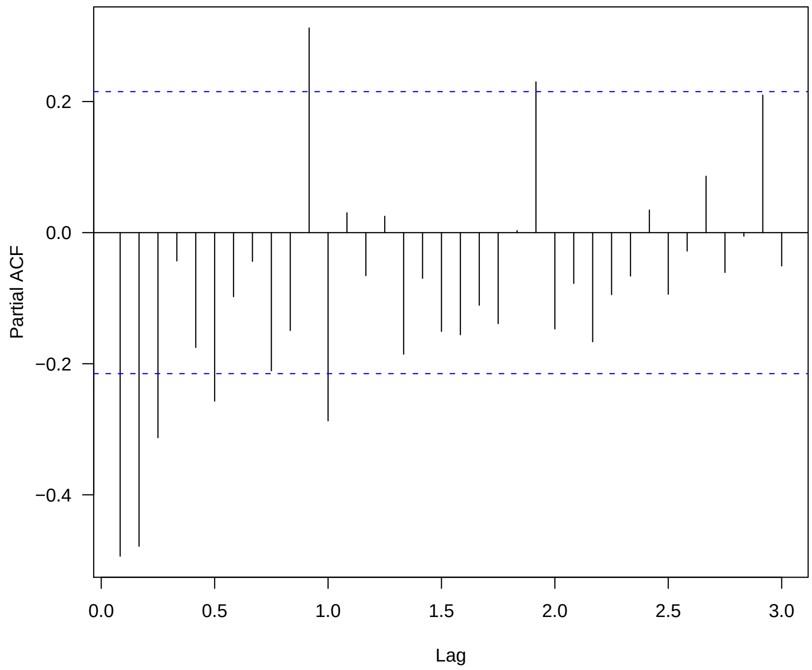

Figure 13: The PACF of the twice-differenced road deaths series.

Table 2: Forecasts for the road deaths series.

Month	Forecast	Std. Error
Jan	29.98	6.31
Feb	30.73	6.31
Mar	32.10	6.32
Apr	31.48	6.32
May	30.10	6.33
Jun	29.10	6.33
Jul	27.35	6.33
Aug	27.10	6.34
Sep	25.60	6.34
Oct	26.85	6.34
Nov	29.35	6.35
Dec	37.85	6.35

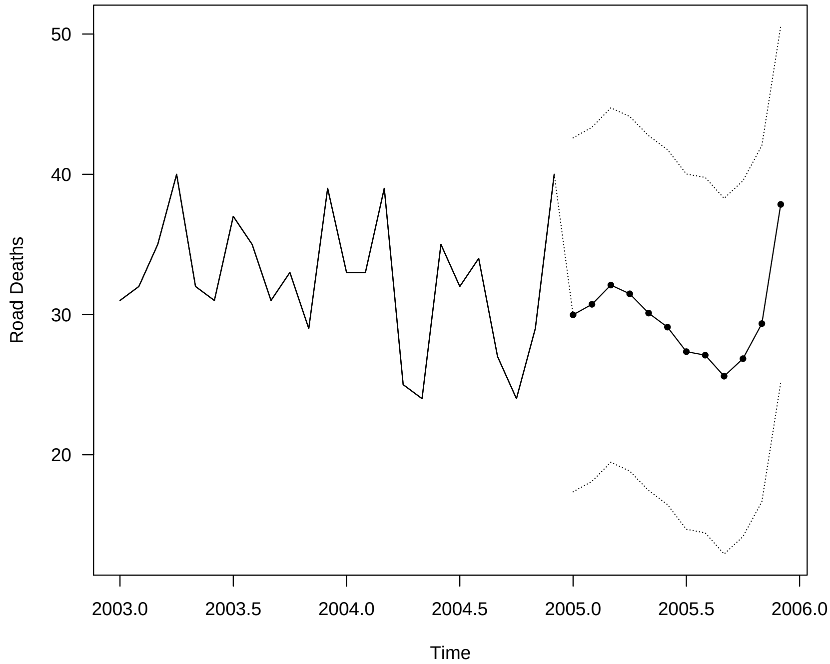

Figure 14: Forecasts for the New Zealand road deaths series.